Hypothesis Testing: One-Sample Inference

7

Solved Problems-Number (2)

One-Sample Inference

Hypothesis Testing for the Population Mean (μ)

7.4 One-Sample Z-Test for the Mean of a Normal Distribution

The test procedure for the one-sample Z test can explained using the three cases of the alternative hypothesis (H_1) as follows:

Test Statistic (Calculated Value)

Case (1)

- Conditions

(1) Normal Distribution.

(2) Population Standard Deviation σ is Known.

(3) Sample size (n) is small ($n < 30$) or is large ($n \geq 30$).

Test Statistic

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

Case (2)

- Conditions

(1) Normal Distribution.

(2) Population Standard Deviation σ is Un Known.

(3) Sample size (n) is large ($n \geq 30$).

Test Statistic

$$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

Notation

If the underlying distribution is unknown or not-normal and the sample size (n) is large ($n \geq 30$) then the central limit theorem can be used as follows:

(a) If population standard deviation σ is known then calculate Z using case (1).

(b) If population standard deviation σ is unknown then calculate Z using case (2).

Rejection Rule

One-Sided Alternative (Lower-Tailed Test)

Hypotheses

$$H_0 : \mu = \mu_0 \ \text{ vs } \ H_1 : \mu < \mu_0$$

Rejection Rule

- Reject H_0 at level of significance α if $Z < -Z_{1-\alpha}$
- Accept H_0 at level of significance α if $Z \geq -Z_{1-\alpha}$

where Z = Test Statistic value and $-Z_{1-\alpha}$ = Critical Value from Table 3 in the Appendix.

One-Sided Alternative (Upper-Tailed Test)

Hypotheses

$$H_0 : \mu = \mu_0 \ \text{ vs } \ H_1 : \mu > \mu_0$$

Rejection Rule

- Reject H_0 at level of significance α if $Z > Z_{1-\alpha}$
- Accept H_0 at level of significance α if $Z \leq Z_{1-\alpha}$

where Z = Test Statistic value and $Z_{1-\alpha}$ = Critical Value from Table 3 in the Appendix.

Two-Sided Alternative (Two-Tailed Test)

Hypotheses

$$H_0 : \mu = \mu_0 \ \text{ vs } \ H_1 : \mu \neq \mu_0$$

Rejection Rule

- Reject H_0 at level of significance α if $|Z| > Z_{1-\left(\frac{\alpha}{2}\right)}$
- Accept H_0 at level of significance α if $|Z| \leq Z_{1-\left(\frac{\alpha}{2}\right)}$

where Z = Test Statistic value and $Z_{1-\left(\frac{\alpha}{2}\right)}$ = Critical Value from Table 3 in the Appendix.

Question (1)

Suppose it is known that in a certain large human population cranial length is normally distributed standard deviation 12.7 mm. A random sample of size 10 is taken from this population showed that the sample mean ($\overline{X}$) is 186.1 mm. Use the one-sample z-test to test the hypothesis:

$$H_0 : \mu = 190 \ \text{ vs } \ H_1 : \mu < 190$$

at level of significance $\alpha = 0.05$?

Solution

- Conditions

(1) Normal Distribution.
(2) Population Standard Deviation σ is known ($\sigma = 12.7$).
(3) Sample size (n) is small ($n = 10 < 30$) (*not important*).

- Test Statistic Value

$$z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{186.1 - 190}{12.7/\sqrt{10}} = -0.971$$

- Rejection Rule (One-Sided Lower Tailed Test)

Reject H_0 at significance level $\alpha = 0.05$ if $Z < -Z_{1-\alpha}$ otherwise Accept H_0.

- Critical Value

$$-Z_{1-\alpha} = -Z_{1-0.05} = -Z_{0.95} = -1.645$$

- Decision

We get $Z = -0.971 > -Z_{0.95} = -1.645$
Then Accept H_0: μ = 190 and therefore Reject H_1: μ < 190 at α= 0.05.

Question (2)

Suppose it is known that in a certain large human population cranial length is normally distributed standard deviation 12.7 mm. A random sample of size 100 is taken from this population showed that the sample mean ($\bar{X}$) is 194.1 mm. Use the one-sample Z-test to test the hypothesis:

$$\mathrm{H}_0 : \mu = 190 \ \text{ vs } \ \mathrm{H}_1 : \mu > 190$$

at level of significance $\alpha = 0.05$?

Solution

- Conditions

(1) Normal Distribution.
(2) Population Standard Deviation σ is known ($\sigma = 12.7$).
(3) Sample size (n) is large ($n = 100 > 30$) (*not important*).

➢ Test Statistic Value

$$z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{194.1 - 190}{12.7/\sqrt{100}} = 3.228$$

➢ Rejection Rule (One-Sided Lower Tailed Test)

Reject H_0 at significance level $\alpha = 0.05$ if $Z > Z_{1-\alpha}$ otherwise Accept H_0.

➢ Critical Value

$$Z_{1-\alpha} = Z_{1-0.05} = Z_{0.95} = 1.645$$

➢ Decision

We get $Z = 3.228 > Z_{0.95} = 1.645$
Then Reject H_0: μ = 190 and therefore Accept H_1: μ < 190 at α= 0.05.

Question (3)

Suppose it is known that in a certain large human population cranial length is normally distributed. A random sample of size 100 is taken from this population showed that the sample mean ($\bar{X}$) is 186.1 mm and the sample standard deviation is 12.7 mm. Use the one-sample Z-test to test the hypothesis:

$$\mathrm{H}_0 : \mu = 190 \quad \text{vs} \quad \mathrm{H}_1 : \mu \neq 190$$

At level of significance $\alpha = 0.05$?

Solution

➢ Conditions

(1) Normal Distribution.
(2) Population Standard Deviation σ is Unknown ($S = 9.7$).
(3) Sample size (n) is large ($n = 100 > 30$).

➢ Test Statistic Value

$$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{186.1 - 190}{12.7/\sqrt{100}} = -3.071$$

➢ Rejection Rule (Two-Sided Tailed)

Reject H_0 at significance level $\alpha = 0.05$ if $|Z| > Z_{1-\alpha/2}$ otherwise Accept H_0.

- Critical Value

$Z_{1-\alpha/2} = Z_{1-0.05/2} = Z_{0.975} = 1.96$

- Decision

We get $|Z| = |-3.071| = 3.071 > Z_{1-\alpha/2} = Z_{0.975} = 1.96$
Then Reject H_0: μ = 190 and therefore Accept H_1: μ ≠ 190 at α= 0.05.

7.5 Sample-Size Determination

EQUATION 7.24

Sample-Size Estimation Based on CI Width

Suppose we wish to estimate the mean of a normal distribution with sample variance s^2 and require that the two-sided 100% × (1 − α) CI for μ be no wider than L. The number of subjects needed is approximately

$$n = 4z^2_{1-\alpha/2}\, s^2/L^2$$

Notation: The value of the sample size (n) obtained by using Equation 7.24 should be rounded up to the next integer.

Question (4)

Cardiology Find the value for the minimum sample size (n) needed to estimate the change in heart rate (μ) if we require that the two-sided 95% CI for μ be no wider than 5 beats per minute and the sample standard deviation for change in heart rate equals 10 beats per minute?

Solution

Step(1)

$(1-\alpha)\,100\% = 95\%$

$1 - \alpha = 0.95$

$\alpha = 0.05$

$\frac{\alpha}{2} = \frac{0.05}{2} = 0.025$

$1 - \frac{\alpha}{2} = 1 - 0.025 = 0.975$

$Z_{1-\frac{\alpha}{2}} = Z_{0.975} = 1.96$

Refer to Table 3 in the Appendix

Step(2)

$$n = \frac{4\left(Z_{1-(\alpha/2)}\right)^2 s^2}{L^2}$$

$$n = \frac{4\,(1.96)^2\,10^2}{5^2}$$

$$n = \frac{(15.3664)(100)}{25}$$

$$n = 61.4656$$

$$n \cong 62$$

7.6 One-Sample Inference for the Binomial Distribution

In this section, we will study the one-sample inference for the population proportion of a binomial distribution (p) based on the sample proportion of cases $\hat{p}$ assuming the normal approximation to the binomial distribution is valid. We know that under $H_0: p = p_0$ and when $np_0q_0 \geq 5$ where $q_0 = 1 - p_0$, then the sampling distribution of the sample proportion ($\hat{p}$) will be normal distribution as follows:

$$\hat{p} \sim N(p_0, \tfrac{p_0 q_0}{n})$$

In the rest of this section, we focus primarily on two-sided tests because they are much more widely used in the literature. Now, to test at level of significance (α) the following two-sided alternative hypotheses: $H_0 : p = p_0$ vs $H_1 : p \neq p_0$. A continuity-corrected version of the test statistic Z can be used. Thus, the test takes the following form:

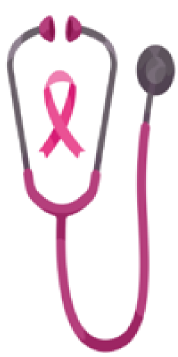
EQUATION 7.27

One-Sample Test for a Binomial Proportion—Normal-Theory Method (Two-Sided Alternative)

Let the test statistic $z_{corr} = \left(|\hat{p} - p_0| - \frac{1}{2n}\right) \Big/ \sqrt{p_0 q_0 / n}$.

If $z_{corr} > z_{1-\alpha/2}$, then H_0 is rejected. If $z_{corr} < z_{1-\alpha/2}$, then H_0 is accepted. This test should only be used if $np_0q_0 \geq 5$.

Question (5)



Cancer Consider a breast-cancer problem were we interested in the effect of having a family history of breast cancer on the incidence of breast cancer. Suppose that 400 of the 10,000 (*that is the value of* $\hat{p} = 400/10000 = 0.04$) women ages (50–54) sampled whose mothers had breast cancer had breast cancer themselves at some time in their lives. Given large studies, assume the prevalence rate of breast cancer for U.S. women in this age group is about 2% ($p = p_0 = 0.02$). If p = prevalence rate of breast cancer in (50-to-54) year-old women whose mothers have had breast cancer, then we want to test the two-sided alternative hypothesis:

$$H_0: p = 0.02 \quad vs \quad H_1: p \neq 0.02$$

at level of significance $\alpha = 0.05$?

Solution

- Test Statistic Value

$$Z_{corr} = \frac{|\hat{p} - p_0| - \frac{1}{2n}}{\sqrt{p_0 q_0 / n}}$$

$$= \frac{|.04 - .02| - \frac{1}{2(10{,}000)}}{\sqrt{.02(.98)/10{,}000}} = \frac{.0200}{.0014} = 14.3$$

Table 3 Appendix

- Critical Value

$Z_{1-\frac{\alpha}{2}} = Z_{0.975} = 1.96$

- Decision

We get

$Z_{corr} = 14.3 > Z_{1-\frac{\alpha}{2}} = Z_{0.975} = 1.96$

It follows that H_0: $p = 0.02$ can be rejected using a two-sided test with α = 0.05 and conclude that $p \neq 0.02$.

p – value

DEFINITION 7.13 The p-**value** for any hypothesis test is the α level at which we would be indifferent between accepting or rejecting H_0 given the sample data at hand. That is, the p-value is the α level at which the given value of the test statistic (such as t) is on the borderline between the acceptance and rejection regions.

p-value Method (Approach)

The p-value Method (Approach) can be used to establish whether results from hypothesis tests are statistically significant:

(1) Calculate the exact p-value.

(2) If p-value $< \alpha$, then H_0 is rejected (*results are statistically significant*).

(3) If p-value $\geq \alpha$, then H_0 is accepted (*results are not statistically significant*).

Example

Suppose that we use the one-sample t test to test the hypothesis:

Ho: μ = 120 vs. H1: μ < 120

based on the birthweight data for $n = 100$ and $\alpha = 0.05$, and after we compute the value of the test statistic we get t = -2.08 and the p-value for the birthweight data is as follows:

$$p = P\left(t_{(n-1)} \leq t\right) = P(t_{99} \leq -2.08) = 0.020$$

Assess the statistical significance of the birthweight data?

Solution

Because the p-value is $p = 0.020 < \alpha = 0.05$, then H_0 is rejected and the results would be considered statistically significant and we would conclude that the true mean birthweight (μ) is significantly lower in this hospital than in the national average in the general population.

Example

Suppose that we use the one-sample t test to test the hypothesis:

Ho: μ = 120 vs. H1: μ < 120

based on the birthweight data for $n = 10$ and $\alpha = 0.05$, and after we compute the value of the test statistic we get t = -1.32 and the p-value for the birthweight data is as follows:

$$p = P\left(t_{(n-1)} \leq t\right) = P(t_9 \leq -1.32) = 0.110$$

Assess the statistical significance of the birthweight data?

Solution

Because the p-value is $p = 0.110 > \alpha = 0.05$, then H_0 is accepted and the results would be considered not statistically significant and we would conclude that the true mean birthweight (μ) does not differ (equal) significantly in this hospital to the national average in the general population.

How to Calculate the p-value for the One-Sample t-Test of the Mean (μ):

(1) For a one-sided lower-tailed t–test:

$$\mathrm{H}_0 : \mu = \mu_0 \ \text{ vs } \ \mathrm{H}_1 : \mu < \mu_0$$

the p-value can be calculated as follows:

$$p = P\big(t_{(n-1)} \le t\big)$$

(2) For a one-sided upper-tailed t–test:

$$\mathrm{H}_0 : \mu = \mu_0 \ \text{ vs } \ \mathrm{H}_1 : \mu > \mu_0$$

the p-value can be calculated as follows:

$$p = P\big(t_{(n-1)} > t\big) = 1 - P\big(t_{(n-1)} \le t\big)$$

(3) For a two-sided (two-tailed) t–test:

$$\mathrm{H}_0 : \mu = \mu_0 \ \text{ vs } \ \mathrm{H}_1 : \mu \neq \mu_0$$

the p-value is computed in two different ways, depending on whether t is less than or greater than 0 as follows:

$$p = \begin{cases} 2 \times P\big(t_{(n-1)} \le t\big) & , \ \ if\ t \le 0 \\ 2 \times \big[1 - P\big(t_{(n-1)} \le t\big)\big] & , \ \ if\ t > 0 \end{cases}$$

How to Calculate the p-value for the One-Sample Z-Test of the Mean (μ):

(1) For a one-sided lower-tailed Z–test:

$$\mathrm{H}_0 : \mu = \mu_0 \ \text{ vs } \ \mathrm{H}_1 : \mu < \mu_0$$

the p-value can be calculated as follows:

$$p = P(Z \le z) = \Phi(z)$$

(2) For a one-sided upper-tailed Z–test:

$$\mathrm{H}_0 : \mu = \mu_0 \ \text{ vs } \ \mathrm{H}_1 : \mu > \mu_0$$

the p-value can be calculated as follows:

$$p = P(Z > z) = 1 - P(Z \le z) = 1 - \Phi(z)$$

(3) For a two-sided (two-tailed) Z–test:

$$\mathrm{H}_0 : \mu = \mu_0 \ \text{ vs } \ \mathrm{H}_1 : \mu \neq \mu_0$$

the p-value is computed in two different ways, depending on whether z is less than or greater than 0 as follows:

$$p = \begin{cases} 2 \times P(Z \le z) = 2 \times \Phi(z) & , \ \ if\ z \le 0 \\ 2 \times [1 - P(Z \le z)] = 2 \times [1 - \Phi(z)] & , \ \ if\ z > 0 \end{cases}$$

Question (6)

Cardiovascular Disease Consider a cholesterol data. Assume that the standard deviation is known to be 40 and the sample size is 200. Assess the significance of the results if we want to test:

$$H_0 : \mu = 190 \text{ vs } H_1 : \mu \neq 190$$

at $\alpha = 0.05$? Where the value of the sample mean is $\bar{X}$ = 181.52?

Solution

- The test statistic is: $Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} = \frac{181.52-190}{40/\sqrt{200}}$ = -2.998
- The p-value is:

  We have Z = -2.998 < 0 then the p-value can be calculated as follows:

  $$\begin{aligned} p = 2 \times P(Z \leq z) = 2 \times \Phi(z) &= 2 \times \Phi(-2.998) \\ &= 2 \times [1- \Phi(2.998)] \\ &= 2 \times [1- \Phi(3.00)] \\ &= 2 \times [1- 0.9987] \\ &= 2 \times 0.0013 \\ &= 0.0026 \approx 0.003 \end{aligned}$$

- Hypothesis Test using the Critical Value Method:

  We get $|Z| = |-2.998| = 2.998 > Z_{1-\left(\frac{\alpha}{2}\right)} = Z_{0.975} = 1.96$

  Then Reject H_0: $\mu = 190$ and Accept H_1: $\mu \neq 190$ at α= 0.05, and conclude that the mean of the population (μ) is different from 190.

- Hypothesis Test using the p-Value Method:

  We get p-value = $0.003 < \alpha = 0.05$ then Reject H_0: $\mu = 190$ and Accept H_1: $\mu \neq 190$ at α= 0.05, and conclude that the population (μ) is significantly different from 190.

How to Calculate the p-value for the One-Sample Z-Test of the Proportion (p):

(3) For a two-sided (two-tailed) Z–test:

$$H_0 : p = p_0 \text{ vs } H_1 : p \neq p_0$$

the p-value is computed in two different ways, depending on whether z is less than or greater than 0 as follows:

$$p = 2 \times [1 - \Phi(Z_{corr})]$$

(Twice the area to the right of Z_{corr} under an N(0, 1) curve)

Question (7)

Given that: $H_0: p = 0.02 \ \ vs \ \ H_1: p \neq 0.02$

Assess the statistical significance of the data at level of significance $\alpha = 0.05$?

Solution

- The test statistic is:

$$Z_{corr} = \frac{|\hat{p} - p_0| - \frac{1}{2n}}{\sqrt{p_0 q_0 / n}}$$

$$= \frac{|.04 - .02| - \frac{1}{2(10,000)}}{\sqrt{.02(.98)/10,000}} = \frac{.0200}{.0014} = 14.3$$

We have $Z_{corr} = 14.3$ then the p-value is:

$$\begin{aligned} p &= 2 \times [1 - \Phi(Z_{corr})] \\ &= 2 \times [1 - \Phi(14.3)] \\ &= 2 \times [1 - 1] \\ &= 2 \times [0] \\ &= 0 < 0.05 \end{aligned}$$

Thus, Reject H0, the results are very highly significant. Then $p \neq 0.02$.

Exercises

Exercise (1)

A certain bag of fertilizer advertises that it contains 7.25 kg, but the amounts these bags actually contain is normally distributed with a mean of 7.4 kg and a standard deviation of 0.15 kg.

The company installed new filling machines, and they wanted to perform a test to see if the mean amount in these bags had changed. Their hypotheses were $H_0: \mu = 7.4$ kg vs. $H_a: \mu \neq 7.4$ kg (where μ is the true mean weight of these bags filled by the new machines).

They took a random sample of 50 bags and observed a sample mean and standard deviation of $\bar{x} = 7.36$ kg and $s_x = 0.12$ kg. They calculated that these results had a P-value of approximately 0.02.

What conclusion should be made using a significance level of $\alpha = 0.05$?

Answer: Since the p-value of 0.02 is less than $\alpha = 0.05$, we should reject H_0 and accept H_a.

Exercise (2)

A particular brand of tires claims that its deluxe tire averages at least 50,000 miles before it needs to be replaced. From past studies of this tire, the standard deviation is known to be 8,000. A survey of owners of that tire design is conducted. From the 28 tires surveyed, the mean lifespan was 46,500 miles with a standard deviation of 9,800 miles. Using alpha = 0.05, is the data highly inconsistent with the claim? Assume normal distribution.

Answer: H_0: $\mu = 50,000$ vs H_1: $\mu < 50,000$

Decision: Reject the null hypothesis.

Exercise (3)

The mean age of Hashemite University students in a previous term was 26.6 years old. An instructor thinks the mean age for online students is older than 26.6. She randomly surveys 56 online students and finds that the sample mean is 29.4 with a standard deviation of 2.1. Conduct a hypothesis test using alpha = 0.05? Assume normal distribution.

Answer: H_0: $\mu = 69{,}110$ vs H_1: $\mu > 69{,}110$

Exercise (4)

A researcher from Jordan claimed that the national unemployment rate is 8%. In a random sample of size 200 residents shows that 22 residents who were unemployed. At $\alpha = 0.05$ and assuming normal distribution, test whether the researcher claim is true or not?

Answer: H_0: $p = 0.08$ vs H_1: $p \neq 0.08$; Do Not Reject the null hypothesis. This means that the researcher claim is true.

Exercise (5)

The mean of the quality point averages of a random sample of 36 college seniors is calculated to be 2.6 and the standard deviation is 0.3. Assume that the distribution is normal? How large a sample is required if we want to be 95% confident that our estimate of µ is off by less than 0.05 (the error is L=0.05)?

Answer: n = 139.
